DOCUMENT RESUME

ED 281 159 CS 008 768

AUTHOR Knafle, June D.; And Others
TITLE Assessing Values in Children's Books.
PUB DATE [85]
NOTE 29p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Childrens Literature; Content Analysis; Elementary Education; *Evaluation Criteria; *Evaluation Methods; Measurement Techniques; *Rating Scales; Reading Material Selection; Reading Research; Supplementary Reading Materials; Teacher Education; Test Validity; *Value Judgment
IDENTIFIERS Caldecott Award

ABSTRACT

A study investigated the presentation of values in children's books, using a scale that was developed specifically for the study. The values categories of the scale were (1) positive behavior, (2) positive feelings, (3) negative behavior, (4) negative feelings, (5) Judeo-Christian religious values, (6) other religious values, (7) traditional values, and (8) neutral values. The scale was applied to every page of the 1938 through 1986 Caldecott winners, and the results were tallied by percentage and by frequency. Decisions as to which values a page presented were based upon a combination of the text and the illustration. The analyses of variance of data from both tallies revealed a significant difference for values, no significant difference for decades, and no significant interaction. A separate analysis of variance of the frequency data revealed positive effects were significant. Newman-Keuls tests showed the percentage tally was more sensitive. The scale could be broadened to include other categories, such as non-traditional values, making it more applicable to contemporary books. The rating scale could be used for a classroom exercise in teacher training courses to focus the students' attention on the values incorporated in children's books, basal readers, and textbooks. (Tables of data and 20 references are included.) (NKA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED281159

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Assessing Values in Children's Books

June D. Knafle

University of Illinois at Chicago

Alice Legenza Wescott

Edmonds, Washington

Ernest T. Pascarella

University of Illinois at Chicago

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

June D. Knafle

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS008768

Abstract

Reports the development of a scale to easily assess values in children's books. The values categories of the scale are Positive Behavior, Positive Feelings, Negative Behavior, Negative Feelings, Traditional Values, Judeo-Christian Religious Values, Other Religious Values, and Neutral Values. The scale was applied to Caldecott winners, 1938-1986, and the results were tallied in two ways, a percentage tally and a frequent tally. Analyses of variance (Values x Decades) were applied, and results are discussed in terms of the differences between the two tallies. Uses of the scale, along with possible adaptations or modifications for research purposes or teacher training classroom exercises are discussed.

Assessing Values in Children's Books

The area of values in children's books is both important and difficult. Its importance lies in the influence that the content of books may have in shaping children's thinking, and its difficulty lies in problems of assessment. Additionally, as society's values continue to change, children's books, as well as adults' books, reflect those changes (Knafle, 1985). Several popular books provide interesting accounts of changes in children's thinking, values, and experiences over the last fifteen years or so (Elkind, 1981; Postman, 1982; Winn, 1983).

Assessing values in children's books has been difficult because of several factors. The main difficulty has been the anecdotal nature of reporting values. A given author who is interested in some specific problem, such as the treatment of racism or sexism in children's books, may often decide to document that problem by citing selected textual materials to support whatever point of view the author holds. While such documentation is often quite informative and helpful, it lacks the objectivity of a controlled study. And there is always the reader's reservation that the author may be finding and citing only what is in agreement with his preconceived views. However, an interesting and quantitative comparison of values in different social studies textbooks is provided by Vitz (1985). But the the majority of the articles do little more than provide guidelines from which to choose books for children. Issues in children's books selection (1973) compiles a number of pertinent articles.

Another problem of assessment is rater agreement. The area of literature, along with children's books and stories, seems especially susceptible to problems of rater agreement, because nuances of meaning may be perceived by one rater and not by another.

The area of moral development is closely aligned to values and seems to have been dominated by the work of Lawrence Kohlberg, as seen by the large number of studies which have used his stages of moral development. Blasi (1980) provides a critical review of related studies along with attention to problems of design, measurement, and interpretation. Kohlberg's stages have even been applied to children's books of various grade levels (Gosa, 1977; Hoskisson & Biskin, 1979; Moir, 1977; Readence, Moore, & Moore, 1982; Stahl-Gemake & Wielen, 1984) and to Newbery winners (Rihn, 1980). However, the validity of Kohlberg's stages has been questioned by a number of researchers (e.g., Kurtines & Greif, 1974; but see Broughton, 1977, for a rebuttal), and it also becomes apparent to anyone who tries to apply his stages to children's books that the difficulties concerning subjectivity are profound, along with the problems of the extensive training time required. Those difficulties may very likely be the result of using the stages in a way in which they were never intended to be used. For a review and discussion of Kohlberg's revised stages, see Nucci and Pascarella (in press).

How to assess values in children's books has thus continued to be a problem primarily because of issues of subjectivity and rater agreement, and this paper reports the development of a scale to easily assess values in children's books.

Evolution of the Scale

When we first decided to assess values in children's books, we assembled a number of descriptions of values from various sources, such as publishers' and reviewers' descriptions, and compiled a list of about a hundred values. We then had about a hundred and fifty graduate and undergraduate students familiarize themselves with the values listed, then individually read children's books and basal reader stories. The students were given a values category sheet for each book and were told to record page numbers of values found in the books on the sheet. For example, next to "cooperation," a student might put pp. 22, 23, 36 from a specific book. Students were told to add values not listed to the sheet. A later variation of that activity was to have page numbers of the book listed and to have students record the number of the value next to the page number. For example, next to page 15, a student might record four or five numbers which corresponded to a master sheet with the values numbered, as well as writing in any values not listed.

There were two main outcomes from these activities. One was the addition of a large number of new categories; a compilation resulted in a total of 364 values, with a number of subcategories.

The other outcome was the realization that there was a high rate of disagreement among the students about the categories of values in the many stories. Many of the disagreements involved shades of meaning not easily reconciled.

Our original intention had been to gather the most specific and exhaustive data possible concerning categories of values and then to collapse the categories into some broad categories. After spending what seemed to be an endless amount of time trying to complete that task, we decided to take a different approach. We examined the 364 categories and decided to group them in different ways. The final categories we decided upon, with examples, are given below.

The Values Categories, with Examples of the Categories

Positive Behavior: cooperation, generosity, helping, honesty, sharing, showing intitiative, showing responsibilty, showing kindness, perseverance, showing friendship

Positive Feelings: patience, liking someone, being happy, positive attitude, courage, self-confidence, feeling empathy

Negative Behavior: aggression, cheating, laziness, lying, stealing, boasting, being nasty, hurting someone on purpose

Negative Feelings; jealousy, hostility, unwarranted anger, being unhappy, selfishness, negative attitude, impatience

Traditional Values: good manners, politeness, greeting each other, neatness, respect for parents, respect for teachers, respect for authority, stereotypical family situations,

patriotism, mention of (or pictures of) traditional celebrations (such as birthday parties, Thanksgiving dinner, Easter egg hunt, wedding), mention of (or pictures of) traditional things (such as Easter bunny, Santa Claus, Christmas tree)

Judeo-Christian Religious Values: trust in God, prayer, praying, forgiveness, mention of (or picture of) priest, nun, minister, rabbi, Jesus, Saint, Christmas, Easter, mass, creche, religious ceremony or festival, church, synagogue, temple

Other Religious Values: mention of (or picture of) witchdoctor, shaman, wizard, religious ceremony or festival, temple, prayer, praying (outside of the Judeo-Christian tradition)

Neutral: dreaming, eating, playing, nature description, surprise, being sick, falling down, giving an incorrect answer, accidently hurting someone, needing privacy, growing up.

We decided not to include such issues as racism and sexism, simply because those issues are well-documented in the literature.

Procedure

We decided to apply the values scale to a discrete set of short books published over a period of time. Winners of the Caldecott medal (one book has been chosen each year since 1938) seemed appropriate. After much thought, we decided to use the individual page of each book as the unit to be measured. For the Caldecott books, every page of each book was measured. The mean number of pages for the Cadecott books was 38. We considered and

rejected the possibilities of using number of sentences or number of words per page or per book for the measurable units as being unworkable for lower level books such as the Caldecott books.

Decisions as to which values a page presented were based upon a combination of the text and picture. Some pages were completely text, some completely picture, or a combination. It did not seem possible or even worthwhile to separate the two because of the illustrative and integrative nature of the two, so both text and pictures were analyzed on each page to determine which values were presented or assumed. Forty-nine books, the Caldecott winners from 1938-1986, were analyzed.

Each book was first analyzed independently by two raters; then the raters discussed and resolved any differences concerning individual values on given pages. (Resolving the differences was not a difficult task once the values criteria had been internalized.)

It also seemed appropriate to analyze the Caldecott winners according to decades in addition to values. We expected to find differences across the decades in Traditional Values, Judeo-Christian Values, Positive Behavior and Feelings, and Negative Behavior and Feelings. These expectations may have been influenced by the view that publishers' standards for appropriate subject matter for children's books have changed over the years, a view easily confirmed by examining current books available for children and comparing those books with children's books available

in the 40's and 50's. The time frames we used for comparing values were 1938-39, the 40's, 50's, 60's, 70's, and 80's.

How the Scale Works

A scale was developed as follows: for each page, a researcher simply put a check mark under the appropriate column. For example, page 1 of a book might have check marks under Positive Feelings and Neutral, page 2 might have check marks under Positive Behavior, Positive Feelings, and Neutral, and page 3 might have check marks under Positive Behavior, Positive Feelings, Negative Feelings, and Neutral.

After the total number of pages had been completed in that way, we decided to tally the results in two different ways: a) a percentage tally, and b) a frequency tally. For the percentage tally, the researcher assigned percentage values for each page. Thus page 1 would have .5 for each of the two values checked, page 2 would have .333 for each of the three values checked, and page 3 would have .25 for each of the four values checked. When all the pages were assigned percentages, the percentages were totaled for each column. The percentage totals were then divided by the total number of pages, giving an overall percentage of each category (column) for the total story. So the overall percentage values for the three-page story would be Positive Behavior, 19.43%; Positive Feelings, 36.10%; Negative Feelings, 8.33%; and Neutral, 36.10% (Total = 99.96%).

For the frequency tally, checks for each value were simply counted. Thus, for the above example the score for Positive Behavior would be 2; Positive Feelings, 3; Negative Feelings, 1; and Neutral, 3. Each score was then divided by the number of pages (in this case, 3) to find the overall score which would be Positive Behavior, .67; Positive Feelings, 1.00; Negative Feelings, .33; and Neutral, 1.00. Table 1 gives an example of a completed story using both tallies. The advisability of having these two different tallies will become clear in the discussion of the findings.

Insert Table 1 about here

Analysis

A two-way analysis of variance using a repeated measures design (Values x Decades) was applied to the percentage tally and to the frequency tally. The Newman-Keuls test was subsequently applied to determine whether pair differences were significant.

The frequency scores for Positive Behavior and Positive Feelings were added together to form the combined value Positive, and the frequency scores for Negative Behavior and Negative Feelings were added together to form the combined value Negative. Positive was then compared with Negative across decades using a two way analysis of variance with repeated measures. The frequency scores were used instead of the percentage scores

because a separate comparison (apart from the other values scores) was desired.

Results

Percentage Tally

The two-way analysis of variance (repeated measures, Values x Decades) revealed a significant difference for Values ($F(7, 301) = 102.19$, $p<.001$), no significant difference for Decades ($F(5, 43) = 1.30$, n.s.), and no significant interaction ($F(35, 301) = .94$, n.s.).

Application of the Newman-Keuls test revealed that for 28 pair comparisons, 12 showed significant differences. Neutral was significantly greater ($p<.001$) than all the other values; Positive Behavior was significantly greater than Negative Behavior ($p<.05$), Judeo-Christian Religious Values ($p<.001$), and Other Religious Values ($p\leq.001$); and Positive Feelings were significantly greater ($p<.05$) then Judeo-Christian Religious Values and Other Religious Values. All other pair comparisons were not significant. Table 2 shows means and standard deviations for all values.

Insert Table 2 about here

Frequency Tally

The two-way analysis of variance (repeated measures, Values x Decades) revealed a significant difference for Values ($F(7, 301) = 85.06$, $p<.001$), no significant difference for Decades ($F(5, 43) =$

1.28, n.s.), and no significant interaction ($F(35, 301) = 1.00$, n.s.).

Application of the Newman-Keuls test revealed that for the 28 pair comparisons, 19 showed significant differences. Neutral was significantly greater ($p < .001$) than all the other values; Positive Behavior was significantly greater than Traditional Values ($p<.05$), Negative Feelings ($p<.05$), Negative Behavior ($p<.001$), Judeo-Christian Religious Values ($p<.001$), and Other Religious Values ($p<.001$); Positive Feelings were significantly greater than Negative Behavior ($p<.05$), Judeo-Christian Religious Values ($p<.001$), and Other Religious Values ($p<.001$); Traditional Values were significantly greater than Judeo-Christian Religious Values ($p<.05$) and Other Religious Values ($p<.005$); Negative Feelings were significantly greater than Judeo-Christian Religious Values ($p<.05$) and Other Religious Values ($p<.01$). All other pair comparisons were not significant. (See Table 2 for means and standard deviations.)

Positive versus Negative

The analysis of variance revealed that the Positive versus Negative effects were significant at the .002 level ($F(1, 43) = 11.47$, Positive mean = .49, s.d. = .28; Negative mean = .29, s.d. = .24), but that the Decade effects were not significant ($F(5, 43) = 1.82$, n.s.).

Graphic Presentation

Although Decade differences were not significant, examination of Figure 1 shows that some interesting information is revealed through graphic presentation, information that would not be available otherwise.

Insert Figure 1 about here

Discussion

The major difference between the two tallies (percentage and frequency), in terms of statistical differences, was that the frequency tally generated more significant pair differences (19 of 28 versus 12 of 28 for the percentage tally). It is always interesting to note the different results obtained by simply looking at data in alternate ways.

Although the statistical results did not differ very much according to the nature of the tally, two outcomes related to the different tallies should be noted. First, the frequency count enabled the researchers to easily combine the values categories (Positive Behavior and Positive Feelings into Positive, Negative Behavior and Negative Feelings into Negative) and delete the other categories so that a statistical analysis could be made of the two main categories of Positive and Negative. Using a frequency tally would seem especially applicable when a researcher considers the possibility of deleting categories from a specific analysis.

Second, being able to compare the results of given text according to the different tallies was useful because the comparison revealed some interesting differences. Table 3 shows several examples of the differences between the tallies for the same books. In Example A, the relationship between Positive Behavior

---

Insert Table 3 about here

---

and Negative Behavior is quite different in the percentage tally (twice as much Negative Behavior) than it is in the frequency tally. In Example B, the relationship between Negative Behavior and Negative Feelings is different in the frequency tally, but not in the percentage tally. In Example C, the relationship of Positive Feelings and Negative Feelings to Neutral is different in the two tallies, with the amount of Positive and Negative Feelings in relation to Neutral appearing to be much greater in the frequency tally than in the percentage tally. In Example D, the relationship of Traditional Values to Neutral is different in the two tallies, with the amount of Traditional Values in relation to Neutral appearing to be much greater in the frequency tally than in the percentage tally.

The differences resulting when the two tallies are compared for any given book are caused by the nature of the tallies. In most instances, the results seem to be approximately comparable. However, in instances when the results are not comparable, it

would seem that the percentage tally, although more cumbersome and time consuming to use than the frequency tally, would reflect a somewhat more sensitive portrayal of the book than the frequency tally. A glance at Table 1 shows that for pages 2 and 3, Negative Feelings are given the weights of .25 and 1.0 for the percentage tally and a count of 2 for the frequency tally. It is quite possible that these frequency differences tend to cancel each other out, especially when a large number of pages is considered. However, a researcher would do well to examine the raw data carefully to be sure that a value appearing only as a small part in conjunction with other values is not equated with values that seem to dominate. That situation could happen when there are a large number of 1.0's for a given value in the percentage tally and an equally large number of .2's (for example) for another value.

The fact that we did not find the major decade differences that we had expected may be accounted for by several factors. It is possible that the award-winning books are showcase books which are not representative of other children's books published during the given years; it has been suggested that the awards' committee tends to be especially conservative. Clearly, Caldecott winners have not been on the cutting edge of contemporary content, especially within the last fifteen years, during which time chidren's book content has increasingly reflected the stark reality of our times and society (Knafle, 1985).

Use of the Scale

The scale, or adaptations thereof, may be used in several ways. For children's books of about fifty pages or less, it would probably be best to apply the scale to every page. For a longer book, after reading the book, the researcher may choose to apply the scale to a given number of pages, say twenty-five, chosen randomly throughout the book. For a book composed of short stories, the researcher may choose several representative stories and apply the scale only to those stories. In choosing short stories, it would be desirable to choose stories from different places in the book, such as the first third, the middle third, and the last third of the book.

The scale may be used as is, or modified according to the purposes of the researcher. We chose not to include such values as racial and sexual stereotypes simply because many researchers have reported on those values, but they could certainly be included. One value variation which probably should be added if the researcher is looking at contemporary books is non-traditional values, since many contemporary books seem to make a point of including non-traditional values, such as disrespect for parents and authority. Such was not the case with the Caldecott winners.

Once in a while, a reader may come across a book that does not lend itself to accurate page by page assessment simply because a character's page by page actions may not reflect his overall motivation. For example, in Master Salt the sailor's son, the book's page by page assessment would be more positve than the book

really is since the positive occurrences are the direct result of the young child disobeying his parents, and a child reader may get the message that good things may happen when parents are disobeyed.

A way to secure a high degree of rater agreement is to make the subjective categories as objective as possible by providing specific examples of the various categories. Even after that is done, there will still be disagreements between raters, so the disagreements will have to be reconciled. We have found that coming to a common agreement is relatively easy by having the raters discuss their differences from specific pages with each other.

The scale could be also used in a classroom exercise in teacher training classes to focus students' attention on the values incorporated in children's books, basal readers, and textbooks. For classroom use, having students work in pairs, discussing their judgments as they read and evaluate each page, seems to work well. Many students have noted that that kind of focused activity made them look at children's stories in a new light.

References

Ahlberg, A. (1980). Master Salt the sailor's son. Racine, WI: Western.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. Psychological Bulletin, 88, 1-45.

Broughton, J. (1977). The cognitive-development approach to morality: A reply to Kurtines and Greif. Journal of Moral Education, 7, 81-95.

Elkind, D. (1981). The hurried child. Reading, MA: Addison-Wesley.

Gosa, C. (1977). Moral development in current fiction for children and young adults. Language Arts, 54, 529-536.

Hoskisson, K., & Biskin, D. S. (1979). Analyzing and discussing children's literature using Kolhberg's stages of moral development. The Reading Teacher, 33, 141-147.

Issues in children's book selection. A School Library Journal/Library Journal Anthology. (1973). New York: R.R. Bowker, Xerox.

Knafle, J. D. (1985). Changing values in children's books. The Reading Instruction Journal, 28(3), 18-21.

Kurtines, W., & Greif, E. B. (1974). The development of moral thought: review and evaluation of Kohlberg's approach. Psychological Bulletin, 81, 453-470.

Moir, H. (1977). If we've always had books that taught all these virtues, why is our society in such lousy shape? Language Arts, 54, 522-528.

Nucci, L., & Pascarella, E. T. (in press). The influence of college on moral development. In P. Terenzini (Ed.) Higher education: Handbook of theory and research, Vol. III. New York: Agathon.

Postman, N. (1982). The disappearance of childhood. New York: Dell.

Readence, J. E., Moore, D. W., & Moore, S. A. (1982). Kohlberg in the classroom: Responding to literature. Journal of Reading, 26, 104-108.

Rihn, B. A. (1980). Kohlberg level of moral reasoning of protagonists in Newbery award winning fiction. Reading Research Quarterly, 15, 377-398.

Stahl-Gemake, J., & Wielan, O. P. (1984). A study guide for building moral reasoning through adolescent literature. Journal of Reading, 28, 34-39.

Vitz, P. C. (1985). Religion and traditional values in public school textbooks: An empirical study. Part of an NIE Grant: Equity in values education: Do the values education apsects of public school curricula deal fairly with diverse belief systems (NIE G-84-0012; Project No. 2-0099).

Winn, M. (1983). Children without childhood. New York: Pantheon (Random House).

Authors

June D. Knafle, Associate Professor, College of Education, University of Illinois at Chicago, Box 4348, Chicago, IL 60680. Specializations: Reading education and children's literature.

Alice Legenza Wescott, 23104 80th Place West, Edmonds, WA 98020. Specialization: Reading education.

Ernest T. Pascarella, Professor, College of Education, University of Illinois at Chicago, Box 4348, Chicago, IL 60680. Specializations: Educational research and evaluation.

Table 1

Completed Story Using the Percentage Tally and the Frequency Tally on the Values Scale

	Positive Behavior	Positive Feelings	Negative Behavior	Negative Feelings	Traditional Values	J-C Religious Values	Other Religious Values	Neutral
Values Scale Page by Page Checklist								
Page No.								
1		✔.5			✔.5			
2	✔.25	✔.25		✔ .25				✔.25
3				✔1.0				
4	✔.333	✔.333						✔.333
5		✔.5						✔.5
Percentage Tally								
Total Score	.583	1.583	--	1.25	.5	--	--	1.083
% Score[a]	11.66	31.66	--	25.0	10.0	--	--	21.66
Frequency Tally								
Total Score	2	4	--	2	1	--	--	3
Frequency Score[a]	.40	.80	--	.40	.20	--	--	.60

[a]Total Score divided by number of pages

Table 2

Means and Standard Deviations for All Values in Caldecott Winners for the Percentage Tally and the Frequency Tally

	Positive Behavior	Positive Feelings	Negative Behavior	Negative Feelings	Traditional Values	J-C Religious Values	Other Religious Values	Neutral
				Percentage Tally				
Mean	11.82	8.72	5.55	6.74	7.17	2.98	2.04	54.96
SD	11.49	6.31	5.84	5.86	7.45	6.78	6.31	16.75
				Frequency Tally				
Mean	.27	.22	.13	.16	.19	.08	.05	.88
SD	.21	.17	.13	.13	.20	.15	.14	.18

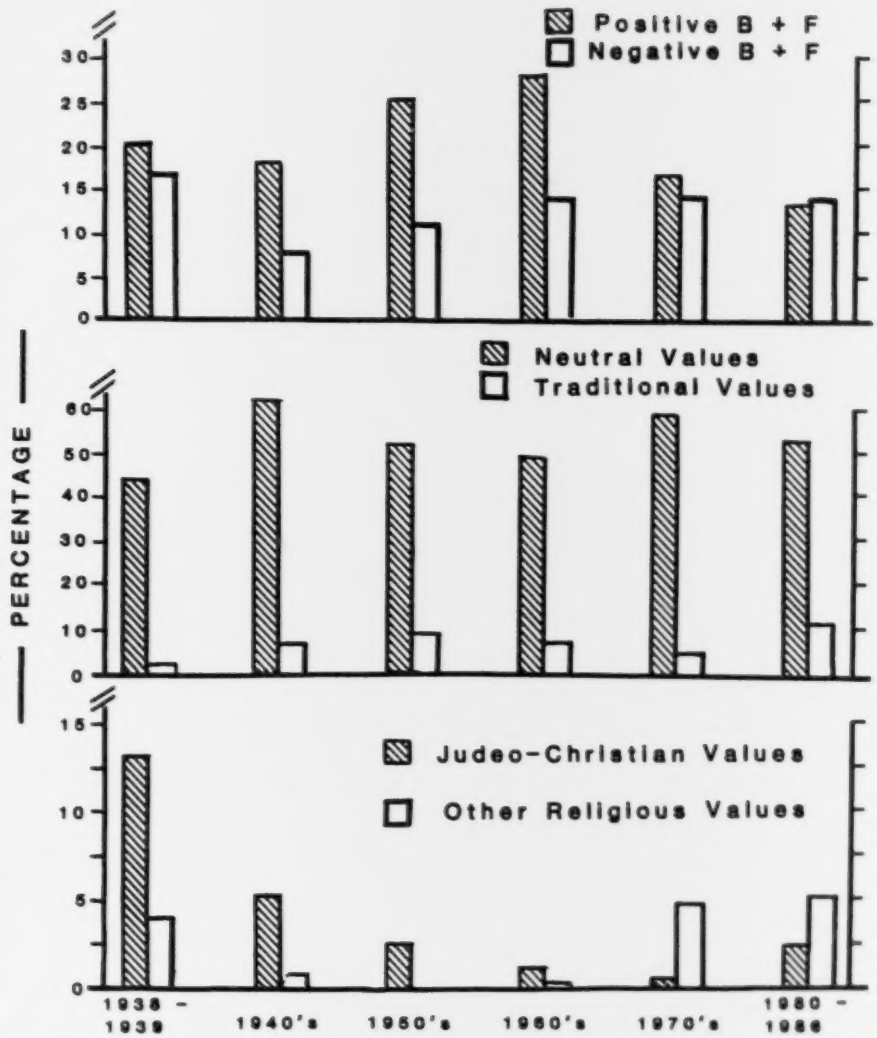
Table 3

Examples of Differences Between the Percentage Tally and the Frequency Tally for Values in Individual Books on the Values Scale

EXAMPLE	Positive Behavior	Positive Feelings	Negative Behavior	Negative Feelings	Traditional Values	J-C Religious Values	Other Religious Values	Neutral
A %	6.66	6.11	12.22	--	22.77	--	--	52.22
A Freq.	.17	.13	.17	--	.43	--	--	.83
B %	12.24	8.72	4.95	4.94	6.12	7.68	--	55.33
B Freq.	.30	.23	.09	.14	.14	.19	--	.88
C %	5.79	7.48	5.79	8.57	4.15	3.61	--	64.58
C Freq.	.16	.21	.16	.25	.13	.10	--	1.00
D %	13.50	1.35	--	--	17.56	--	--	67.56
D Freq.	.35	.03	--	--	.43	--	--	1.00

Figure Caption

Figure 1. Percentage of values in Caldecott winners according to decades.


Positive B + F
Negative B + F
Neutral Values
Traditional Values
Judeo-Christian Values
Other Religious Values
PERCENTAGE
1938 - 1939
1940's
1950's
1960's
1970's
1980 - 1988